

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 25, 2005, 11:42:31 ; Search time 161.856 Seconds
(without alignments)
4214.186 Million cell updates/sec

Title: US-09-982-091a-4
Perfect score: 6836
Sequence: 1 MTGEVGEVHLEINDPNVIS.....LTWSGAPRGFRLSPDPHG 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6836	100.0	1332	2 Q9HAW4	Q9HAW4 homo sapien
2	4838	70.8	1315	2 Q80YR7	Q80YR7 mus musculu
3	3512	51.4	693	2 Q8IWI1	Q8IWI1 homo sapien
4	2687	39.3	1285	2 Q9DP50	Q9DP50 xenopus lae
5	2340	34.2	462	2 Q9BX02	Q9BX02 homo sapien
6	1618.5	23.7	411	2 Q95GM2	Q95GM2 mus musculu
7	1500	21.9	289	2 Q9BX01	Q9BX01 homo sapien
8	836	12.2	165	2 Q9BX03	Q9BX03 homo sapien
9	784	11.5	270	2 Q8BPX6	Q8BPX6 mus musculu
10	700	10.2	1465	2 Q8IRB5	Q8IRB5 drosophila
11	534.5	7.8	180	2 Q8BRD4	Q8BRD4 mus musculu
12	526	7.7	835	2 Q7PR25	Q7PR25 anopheles g
13	497	7.3	105	2 Q6PEH5	Q6PEH5 homo sapien
14	487.5	7.1	7210	2 Q9V7G8	Q9V7G8 drosophila
15	487.5	7.1	9270	2 Q8MLD9	Q8MLD9 drosophila
16	454	6.6	2081	2 Q9LH98	Q9LH98 arabisdopsis
17	448.5	6.6	725	2 Q17838	Q17838 caenorhabdi
18	448	6.6	701	2 Q8MYR6	Q8MYR6 drosophila
19	446.5	6.5	3404	2 Q7RJB3	Q7RJB3 plasmodium
20	419	6.1	1464	2 Q871F7	Q871F7 neurospora
21	415	6.1	1661	2 Q06166	Q06166 plasmodium
22	413.5	6.0	10578	2 Q8ISF5	Q8ISF5 caenorhabdi
23	413.5	6.0	18519	2 Q8ISF6	Q8ISF6 caenorhabdi
24	413.5	6.0	18534	2 Q8ISF7	Q8ISF7 caenorhabdi
25	411	6.0	2301	2 Q6BG00	Q6BG00 paramedum
26	410	6.0	1451	2 Q812D8	Q812D8 plasmodium
27	409	6.0	5507	2 Q8IHN3	Q8IHN3 plasmodium
28	407.5	6.0	5458	2 Q9U4S9	Q9U4S9 plasmodium
29	402.5	5.9	1481	2 Q8EVB9	Q8EVB9 mycoplasma
30	396.5	5.8	1510	2 Q25920	Q25920 plasmodium
31	395.5	5.8	1871	2 Q9SRD5	Q9SRD5 arabisdopsis

32	395.5	5.8	2760	2 Q8ISY2	Q8ISY2 plasmodium
33	394	5.8	5412	1 FUTS DROME	Q9W56 drosophila
34	391	5.7	791	2 Q9DGL1	Q9DGL1 fuigu rubrip
35	391	5.7	1434	2 Q81492	Q81492 plasmodium
36	388.5	5.7	2468	1 MABP_HUMAN	P46821 homo sapien
37	385	5.6	3504	2 Q8IL45	Q8IL45 plasmodium
38	384.5	5.6	2116	1 MYSD2 DICDI	P08799 dictyostel
39	383.5	5.6	1431	2 Q96723	Q96723 homo sapien
40	381	5.6	2464	2 Q64ID6	Q64ID6 macropus eu
41	381	5.6	2910	2 Q9FND5	Q9FND5 arabisdopsis
42	380	5.6	1989	2 Q814U7	Q814U7 plasmodium
43	380	5.6	2095	2 Q7RN75	Q7RN75 plasmodium
44	376	5.5	1805	1 NEST RAT	P21263 ratus norv
45	375	5.5	2459	1 MABP_RAT	P15205 ratus norv

ALIGNMENTS

RESULT 1	Q9HAW4	PRELIMINARY	PRT	1332 AA.
AC	Q9HAW4			
DT	01-MAR-2001 (Tremblrel. 16, Created)			
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Hu-ClaspIn.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21000493; PubMed=11090622;			
RA	Kumagai A., Dunphy W.G.;			
RT	"ClaspIn, a novel protein required for the activation of Chk1 during a			
RT	DNA replication checkpoint response in Xenopus egg extracts."			
RL	MOL. Cell 6:839-849(2000).			
DR	EMBL: AF297866; AAC24515.1; -			
DR	GeneW, HGNC:19715; CUSPN.			
GO	GO:000515; F:Protein binding; NAS.			
SQ	SEQUENCE 1332 AA; 150174 MW; 4EA38DC16A3936C3 CRC64;			
Query Match	100.0%; Score 6836; DB 2; Length 1332;			
Best Local Similarity	100.0%; Pred. No. 8.3e-227;			
Matches 1332; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MTGEVGEVHLEINDPNVISQEBADSPDSGQSYETIGPLSGDSDEIFVSKLKNRK 60			
DB	1 MTGEVGEVHLEINDPNVISQEBADSPDSGQSYETIGPLSGDSDEIFVSKLKNRK 60			
QY	61 VIQDSSETEEDNASEKTTYSAEENKENTLYAGKNTYIKRIYKTVVADSDESYMEKSLY 120			
DB	61 VIQDSSETEEDNASEKTTYSAEENKENTLYAGKNTYIKRIYKTVVADSDESYMEKSLY 120			
QY	121 QENLEAVQVPCLELSQSGNSTPTFTDRKSSKXIHDKGTGKAIVKSKRRLKEKERKM 180			
DB	121 QENLEAVQVPCLELSQSGNSTPTFTDRKSSKXIHDKGTGKAIVKSKRRLKEKERKM 180			
QY	181 EKIRQLKKKETNQEDVDQPPNDGCLLVKDLFETGLEDENNSPLEDEESLESIRAAV 240			
DB	181 EKIRQLKKKETNQEDVDQPPNDGCLLVKDLFETGLEDENNSPLEDEESLESIRAAV 240			
QY	241 KKKVKKKKKKPSLSGCVHSFEESGLSGTTRKEKKAARLSKEALKQHSQRLIRRS 300			
DB	241 KKKVKKKKKKPSLSGCVHSFEESGLSGTTRKEKKAARLSKEALKQHSQRLIRRS 300			
QY	301 AANLPYHMPENKTIHDFGRKRPPTCHGNAMALLKSSKYQSSHHKEIIDTANTTENNSDH 360			
DB	301 AANLPYHMPENKTIHDFGRKRPPTCHGNAMALLKSSKYQSSHHKEIIDTANTTENNSDH 360			
QY	361 HSKGSEQTGAENEVTNALPVVSKETQIIITGSDSCRKDLVKNLELEIOEKOKOSDIRP 420			
DB	361 HSKGSEQTGAENEVTNALPVVSKETQIIITGSDSCRKDLVKNLELEIOEKOKOSDIRP 420			

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:27:19 : Search time 34.1016 Seconds
(without alignments)
2915.770 Million cell updates/sec

Title: US-09-982-091a-4
Perfect score: 6836
Sequence: 1 MTGGEVSEVHLINDEPNVIS.....LTWGAIPGPFRLSPDPHG 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5555	81.3	1089	4	US-09-949-016-10326
2	388.5	5.7	2468	4	US-09-976-594-726
3	388.5	5.7	2468	4	US-09-538-092-1135
4	388.5	5.7	2522	4	US-09-949-016-10237
5	382.5	5.6	1400	4	US-09-764-176-7
6	376	5.5	1805	1	US-07-853-913-2
7	352.5	5.2	2375	4	US-09-538-092-1131
8	349	5.1	1898	2	US-08-800-644-94
9	349	5.1	1898	2	US-08-800-644-94
10	349	5.1	1898	2	US-09-538-092-1280
11	348	5.1	2359	4	US-09-849-016-6507
12	341	5.0	2733	4	US-09-949-016-11433
13	333.5	4.9	3678	4	US-09-814-259-11
14	329.5	4.8	2476	4	US-09-824-574-7
15	327	4.8	2954	4	US-09-150-867-1
16	323.5	4.7	8991	4	US-08-714-741-32
17	313.5	4.6	1939	3	US-09-310-187A-1
18	312.5	4.6	1935	4	US-09-538-092-916
19	312	4.6	1786	3	US-08-973-462-8
20	311.5	4.5	1944	4	US-09-949-016-10929
21	310.5	4.5	1939	4	US-09-538-092-917
22	308	4.5	2663	4	US-09-538-092-1252
23	304.5	4.5	1886	3	US-08-938-105-3
24	302	4.4	3686	4	US-09-134-001C-5080
25	301	4.4	1857	4	US-09-917-254-91
26	301	4.4	1972	4	US-09-538-092-1084
27	301	4.4	1976	4	US-09-538-092-1078

28	301	4.4	1984	4	US-09-949-016-7111	Sequence 7111, Ap
29	301	4.4	1984	4	US-09-949-016-7112	Sequence 7112, Ap
30	301	4.4	1984	4	US-09-949-016-7113	Sequence 7113, Ap
31	300.5	4.4	1427	4	US-09-538-092-1044	Sequence 1044, Ap
32	300	4.4	3924	4	US-09-538-092-1246	Sequence 1246, Ap
33	299.5	4.4	1942	4	US-09-949-016-8135	Sequence 8135, Ap
34	299.5	4.4	2662	4	US-09-595-684B-31	Sequence 31, Appl
35	298	4.4	1618	1	US-07-853-913-4	Sequence 4, Appl
36	298	4.4	1618	4	US-09-538-092-1143	Sequence 1143, Ap
37	297.5	4.4	1572	4	US-08-875-435B-3	Sequence 3, Appl
38	297	4.3	1972	4	US-09-418-710-21	Sequence 21, Appl
39	297	4.3	1972	4	US-09-839-479-21	Sequence 21, Appl
40	296.5	4.3	1969	4	US-09-418-710-72	Sequence 72, Appl
41	296.5	4.3	1969	4	US-09-839-479-71	Sequence 71, Appl
42	296	4.3	1085	1	US-08-431-080-28	Sequence 28, Appl
43	296	4.3	1085	2	US-08-938-534-28	Sequence 28, Appl
44	296	4.3	1085	3	US-09-345-294-28	Sequence 28, Appl
45	296	4.3	2349	4	US-09-538-092-914	Sequence 914, Appl

ALIGNMENTS

```
RESULT 1
US-09-949-016-10326
Sequence 10326, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10326
LENGTH: 1089
TYPE: PRT
ORGANISM: Human
US-09-949-016-10326:
Query Match      81.3%: Score 5555; DB 4; Length 1089;
Best Local Similarity 99.6%: Pred. NO. 0;
Matches 1081; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY      248 KKKESLSGVSHPGSELSKGTTRKKAARLSKALKOLHSETORLIRSALNLPYH 307
      :|||||
DB      5 OKKESLSGVSHPGSELSKGTTRKKAARLSKALKOLHSETORLIRSALNLPYH 64
QY      308 MEENTIHDFPKRRPPTCHGNAMLLSKSKQSSHHKEIITANTTENSDDHSGSQ 367
      :|||||
DB      65 MEENTIHDFPKRRPPTCHGNAMLLSKSKQSSHHKEIITANTTENSDDHSGSQ 124
QY      368 TTGAENEVEYNALPVVSKETQIITGSDSCRDOLVKNELETOEKOKOSDIPSPDSSV 427
      :|||||
DB      125 TTGAENEVEYNALPVVSKETQIITGSDSCRDOLVKNELETOEKOKOSDIPSPDSSV 184
QY      428 LQESNFIQNNHSECOVGLVAFEPHALLEGEPONPEBTDEKVEEPEOONKSSAVGPE 487
      :|||||
DB      185 LQESNFIQNNHSECOVGLVAFEPHALLEGEPONPEBTDEKVEEPEOONKSSAVGPE 244
QY      488 KRRRTLDRLKQLGVDVSIKPRLGADSDSFVLEPETRRELALQRFKKNAPPAKPPA 547
      :|||||
DB      245 KRRRTLDRLKQLGVDVSIKPRLGADSDSFVLEPETRRELALQRFKKNAPPAKPPA 304
QY      548 GGTAVNVNVIKDMGTGDKELKADVVPVTLAPKKLDGASHTRPKBLQVTKAKLQEAMKL 607
      :|||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:37:55 ; Search time 146.586 Seconds
(without alignments)
3793.838 Million cell updates/sec

Title: US-09-982-091a-4

Perfect score: 6836

Sequence: 1 MTGEVSEVHLLEINDPNVIS.....LTWGAIPGFRLSPDPHG 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1862994 seqs, 417510619 residues

Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/2/pubppaa/PCIT_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppaa/PCITUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
12: /cgn2_6/prodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
17: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
18: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/prodata/2/pubppaa/US11_PUBCOMB.pep:*
20: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep:*
21: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep:*
22: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	6836	100.0	US-09-982-091a-4	Sequence 4, Appl1
2	2687	39.3	US-09-982-091a-2	Sequence 2, Appl1
3	850	12.4	US-10-106-698-5533	Sequence 5533, Ap
4	697	10.2	US-11-097-143-858	Sequence 858, App
5	388.5	5.7	US-10-755-889-615	Sequence 615, App
6	388.5	5.7	US-10-489-740-216	Sequence 216, App
7	388.5	5.7	US-10-450-763-46955	Sequence 46955, A
8	382.5	5.6	US-10-094-466-30	Sequence 30, Appl1
9	382.5	5.6	US-09-764-116-7	Sequence 7, Appl1
10	375	5.5	US-10-205-331-66	Sequence 66, Appl
11	365.5	5.3	US-10-732-923-3341	Sequence 3341, Ap

12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
359.5	358.5	354	353.5	353.5	353.5	353.5	353.5	353.5	353.5	353.5	352.5	352.5	352	351	350	348	347.5	347.5	347	346.5	346	343.5	343.5	343.5	342	341	341	341	337.5	333.5	333.5	333	
5.3	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.9	4.9	4.9	4.9	
1679	1234	2758	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	2288	
US-10-369-493-22080	US-10-741-849-7132	US-10-097-143-31983	US-10-732-923-8885	US-10-934-998-100	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888	US-10-732-923-8888

ALIGNMENTS

RESULT 1
US-09-982-091a-4
Sequence 4, Application US/09982091A
Patent No. US20020151030A1
GENERAL INFORMATION:
APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
APPLICANT: KUMAGAI, Akiko
APPLICANT: DUNPHY, William
TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
FILE REFERENCE: CITI1320-1
CURRENT FILING DATE: 2002-10-17
PRIOR APPLICATION NUMBER: US 60/241,246
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1332
TYPE: PRT
ORGANISM: Homo sapiens
US-09-982-091a-4

Query Match 100.0%; Score 6836; DB 9; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTGEVSEVHLLEINDPNVISQEEADSPDSGQGSYETIGPLSGDSDEEIFVSKKLNKR 60
DB 1 MTGEVSEVHLLEINDPNVISQEEADSPDSGQGSYETIGPLSGDSDEEIFVSKKLNKR 60
QY 61 VLDDSEETDTNASEKTTTYSAEENKENLYAGKVTIKRKIRYKTVVADSESYMSKLY 120
DB 61 VLDDSEETDTNASEKTTTYSAEENKENLYAGKVTIKRKIRYKTVVADSESYMSKLY 120

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:53:46 / Search time 37.6645 Seconds
(without alignments)
3402.693 Million cell updates/sec

Title: US-09-982-091A-4
Perfect score: 6836
Sequence: 1 MTGCVSEVHLINDEPNVIS.....LTWSGAPPGFRLSPDPHG 1332

Scoring table: BLOSUM62
Gapop 10.0, Gapexc 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	448.5	6.6	725	2 T21363	hypothetical prote
2	400	5.9	3488	2 T34418	hypothetical prote
3	396.5	5.8	1526	2 A45605	mature-parasite-in
4	395.5	5.8	1871	2 D96796	probable heat choc
5	384.5	5.6	2116	2 A26655	myosin heavy chain
6	379	5.5	5337	2 T13564	microtubule-associ
7	376	5.5	1805	2 A34736	nestin - rat
8	375	5.5	2364	2 A56577	microtubule-associ
9	371.5	5.4	2454	1 QRMSP1	microtubule-associ
10	366.5	5.4	2139	2 T18296	myosin heavy chain
11	365.5	5.3	1909	2 A45592	liver stage antige
12	359.5	5.3	1679	2 S48385	hypothetical prote
13	359	5.3	678	2 A54514	glutamic acid-rich
14	356.5	5.2	1233	2 S56271	hypothetical prote
15	351.5	5.1	1110	2 IS1116	NF-180 - sea lamp
16	350	5.1	1407	1 S28589	trichohyalin - hu
17	349	5.1	1898	1 A45973	trichohyalin - hu
18	348	5.1	3259	1 A56539	giantin - human
19	341	5.0	1325	2 A52300	giantin - human
20	337.5	4.9	1939	2 T18372	repetit organellar
21	337	4.9	1979	2 C17632	hypothetical prote
22	334.5	4.9	1922	2 T00637	hypothetical prote
23	331	4.8	1269	2 F84730	probable myosin he
24	330.5	4.8	1019	2 T50251	hypothetical coile
25	330	4.8	1738	2 T14867	interapfin - slime
26	328.5	4.8	1790	2 S67593	transpore protein
27	327.5	4.8	1992	2 A47297	myosin heavy chain
28	327	4.8	2954	2 T14156	kinesin-related pr
29	326.5	4.8	1549	1 A40691	trichohyalin - she

30	325.5	4.8	1271	2 A45555	glutamate rich pro
31	325.5	4.8	3187	2 JCS637	364k Golgi complex
32	324.5	4.7	1744	2 JH0720	canbin - African
33	324	4.7	2526	2 T20531	hypothetical prote
34	323.5	4.7	1658	2 S55101	hypothetical prote
35	323.5	4.7	2020	2 T21174	hypothetical prote
36	322.5	4.7	1931	2 A59234	slow myosin heavy
37	321	4.7	1804	2 T34518	nestin - golden ha
38	320	4.7	1133	2 T22976	hypothetical prote
39	319.5	4.7	1934	2 T48153	myosin heavy chain
40	318.5	4.7	1935	2 A59286	myosin heavy chain
41	318	4.7	2253	2 T30336	nuclear/mitotic ap
42	317.5	4.6	1432	2 B85431	trichohyalin like
43	317.5	4.6	1957	2 T38077	hypothetical coile
44	316.5	4.6	1192	2 A71623	probable secreted
45	315.5	4.6	1935	1 S06006	myosin beta heavy

ALIGNMENTS

RESULT 1

T21363 hypothetical protein F25H5.5 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004

C/Accession: T21363

R/Steward, C. Submitted to the EMBL Data Library, October 1996

A/Reference number: Z19412

A/Accession: T21363

A/Status: Preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1725 <WIL>

A/Cross-references: UNIPROT:O17938; EMBL:Z01068; PDB:CB02986.1; GSPDB:GN00019; CESP:F2

A/Experimental source: Clone F25H5

C/Genetics:

A/Gene: CESP:F25H5.5

A/Map position: 1

A/Intons: 5/3; 106/1; 129/1; 233/2; 294/2; 529/2; 601/3

Query Match 6.6%; Score 448.5; DB 2; Length 725;
Best Local Similarity 24.1%; Pred. No. 2e-10;
Matches 216; Conservative 152; Mismatches 282; Indels 247; Gaps 40;

QY	466	ETDEKVEEPEQNKSSAVGPEPEKVRFTLDRLKQGVDS--IKRPLGADDEDFVILEPE	523
DB	15	EVDKQAE-----DRKLR--LNRKVALERSPLVPKLR---VILSAN	55
QY	524	-----TNRELEALKQRFWMGHANPAKPRAGQTVNVAVIVKMGTDGKEELKADVFTL-	577
DB	56	STTDLTNEDDDPNKKFKFTFNAEKTPKLAKS-----PEEKLFVAVRKIS	101
QY	578	-APKLLDGAHTKPEKQLVTLKAKIQEAMKLRFRFERQKQALFKLND--GSEEEE	632
DB	102	KNAFKSLTSS-----ETLKKKQASMSAKRRKEQQRKMYEDENHLKAGSDEEE	152
QY	633	EEEEEMTDSBEDGEKVEKEKEEELKEEEEGEESEEGNQTAFLLSSSEITKDE	692
DB	153	EEDVVF-----KKKGGKRVNEEYSEDEDE-----YEPENSG	187
QY	693	KEMDKENNDSSEIGKAVGFLSVPKSLSDSTLLFLKDSSSKMGVSPTEEKSETDENSG	752
DB	188	EDSDKKN-----GSA-----	198
QY	753	QPEKLDDEDSCLLTKESSHNSFELIGSTIPYPCN--RGTGRTSFPFPAAGGRSPS	810
DB	199	-----DDMSVVL-----NDSF-----TYDVFNVVKSSGGS-----TVSNDSRS	235
QY	811	PGIFRASIVASAKSSGKLEPSLP--IEDSODIVNASPEPTLFLGAG--DFOGCLRD	865
DB	236	---EADV---EKNGVQLPTMTETITGSSNDILNLGSGFVGVSQACFPDTLQMLIE	287

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM proteoin - protein search, using sw model

Run on: October 25, 2005, 11:37:55 ; Search time 141.414 Seconds
(without alignments)
3793.838 Million cell updates/sec

Title: US-09-982-091A-2

Perfect score: 6604
Sequence: 1 MALCEEQVFLPEPDISIK.....RFRKSTPVKRSIFOLLE 1285

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1862994 seqs, 417510619 residues

Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09A_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11A_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US11A_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6604	100.0	1285	9	US-09-982-091A-2
2	2687	40.7	1332	20	US-09-982-091A-4
3	784.5	11.9	1812	20	US-11-097-143-858
4	391	5.9	2454	17	US-10-732-923-8884
5	390.5	5.9	6815	20	US-11-097-143-27225
6	378.5	5.7	3257	20	US-11-097-143-29298
7	373	5.6	2492	17	US-10-697-525-2
8	373	5.6	2492	16	US-10-732-923-8883
9	372	5.6	2375	16	US-10-408-765A-277
10	372	5.6	2375	17	US-10-732-923-8881
11	372	5.6	2492	17	US-10-732-923-8882

12	372	5.6	2492	18	US-10-934-998-76	Sequence 76, App1
13	371	5.6	2288	17	US-10-732-923-8885	Sequence 8885, Ap
14	371	5.6	2288	18	US-10-934-998-100	Sequence 100, App
15	371	5.6	2237	17	US-10-732-923-8888	Sequence 8888, Ap
16	371	5.6	2375	17	US-10-732-923-8880	Sequence 8880, Ap
17	365.5	5.5	1234	17	US-10-741-849-7132	Sequence 7132, Ap
18	354	5.4	1398	15	US-10-094-466-10	Sequence 30, App1
19	354	5.4	1400	9	US-09-764-176-7	Sequence 7, App1
20	341.5	5.2	1909	17	US-10-732-923-3341	Sequence 3341, Ap
21	341	5.2	2297	20	US-11-097-143-3334	Sequence 2334, Ap
22	338.5	5.1	1444	16	US-10-437-963-150334	Sequence 150334, Ap
23	331.5	5.0	3111	20	US-11-097-143-7773	Sequence 7773, Ap
24	329.5	5.0	2476	10	US-09-824-574-7	Sequence 7, App1
25	329.5	5.0	2476	17	US-10-732-923-8136	Sequence 8136, Ap
26	324	4.9	1881	17	US-10-033-585-7646	Sequence 7646, Ap
27	323.5	4.9	1875	17	US-10-732-923-3334	Sequence 3334, Ap
28	321.5	4.9	1015	17	US-10-732-923-3329	Sequence 3329, Ap
29	321.5	4.9	3225	16	US-10-408-765A-254	Sequence 254, App
30	320.5	4.9	1875	15	US-10-369-493-42285	Sequence 2285, A
31	320.5	4.9	1875	17	US-10-732-923-3335	Sequence 3335, Ap
32	318	4.8	1690	20	US-11-097-143-10224	Sequence 10224, A
33	318	4.8	1690	20	US-11-097-143-10311	Sequence 10311, A
34	317.5	4.8	2020	15	US-10-369-493-5128	Sequence 5128, Ap
35	317.5	4.8	2020	15	US-10-369-493-5129	Sequence 5129, Ap
36	317	4.8	2910	17	US-10-732-923-3342	Sequence 3342, Ap
37	315.5	4.8	1268	17	US-10-732-923-3332	Sequence 3332, Ap
38	315.5	4.8	1790	15	US-10-369-493-1586	Sequence 1586, Ap
39	315.5	4.8	1790	17	US-10-732-923-3330	Sequence 3330, Ap
40	315.5	4.8	1790	17	US-10-732-923-3331	Sequence 3331, Ap
41	315.5	4.8	2768	20	US-11-097-143-31983	Sequence 31983, A
42	314.5	4.8	3699	14	US-10-171-311-4	Sequence 4, App1
43	314.5	4.8	3317	14	US-10-171-311-8	Sequence 8, App1
44	310.5	4.7	3307	14	US-10-171-311-2	Sequence 2, App1
45	310.5	4.7	3308	18	US-10-756-149-5739	Sequence 5739, Ap

ALIGNMENTS

RESULT 1
US-09-982-091A-2
; Sequence 2, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, AKIKO
; APPLICANT: DUNPHY, WILLIAM
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CITI320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; PRIOR FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1285
; TYPE: PRT
; ORGANISM: Xenopus laevis
; US-09-982-091A-2

Query Match 100.0%; Score 6604; DB 9; Length 1285;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALCEEQVFLPEPDISIKIVETSDSGGSCENADONKLLGVCEDKDTDEILVRKKS 60
DB 1 MALCEEQVFLPEPDISIKIVETSDSGGSCENADONKLLGVCEDKDTDEILVRKKS 60
QY KKEVYVSDSDEELMKNFADNVKSHDNEENETMAVREPKRISAVVDSNDSHE 120
DB KKEVYVSDSDEELMKNFADNVKSHDNEENETMAVREPKRISAVVDSNDSHE 120

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:27:19 ; Search time 32.8984 Seconds
(without alignments)
2915.770 Million cell updates/sec

Title: US-09-982-091a-2

Perfect score: 6604
Sequence: 1 MALCEEQVLEPEDIISL.....RFRDSTPVKRSIFOLLE 1285

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	36.7	1089	US-09-949-016-10326	Sequence 10326, A
2	372	5.6	2375	US-09-538-092-1131	Sequence 1131, Ap
3	354	5.4	1400	US-09-764-176-7	Sequence 7, Appl
4	336.5	5.1	1898	US-08-056-200-94	Sequence 94, Appl
5	336.5	5.1	1898	US-08-800-644-94	Sequence 1280, Ap
6	329.5	5.0	2476	US-09-824-574-7	Sequence 7, Appl
7	321.5	4.9	3259	US-09-949-016-11433	Sequence 11433, A
8	321.5	4.7	3878	US-09-949-016-6507	Sequence 6507, Ap
9	307.5	4.7	3878	US-09-949-016-6507	Sequence 11, Appl
10	303.5	4.6	1786	US-08-973-462-8	Sequence 8, Appl
11	299.5	4.5	2954	US-09-150-867-1	Sequence 1, Appl
12	296.5	4.5	1886	US-08-578-277A-4	Sequence 1, Appl
13	286	4.5	8991	US-08-714-741-32	Sequence 32, Appl
14	293	4.4	2662	US-09-595-684B-31	Sequence 31, Appl
15	293	4.4	2663	US-09-538-092-1252	Sequence 1252, Ap
16	292	4.4	2047	US-09-949-016-7404	Sequence 7404, Ap
17	292	4.4	1857	US-09-917-254-91	Sequence 91, Appl
18	291	4.4	1972	US-09-538-092-1084	Sequence 1084, Ap
19	291	4.4	1979	US-09-949-016-6468	Sequence 6468, Ap
20	291	4.4	1984	US-09-949-016-7111	Sequence 7111, Ap
21	291	4.4	1984	US-09-949-016-7112	Sequence 7112, Ap
22	291	4.4	1984	US-09-949-016-7113	Sequence 7113, Ap
23	290.5	4.4	1960	US-09-538-092-1077	Sequence 1077, Ap
24	290.5	4.4	1960	US-09-949-016-10872	Sequence 10872, A
25	290.5	4.4	1960	US-09-538-092-901	Sequence 901, Appl
26	290	4.4	1963	US-09-949-016-8888	Sequence 8888, Ap
27	290	4.4	1963	US-09-949-016-8888	Sequence 8888, Ap

28	288.5	4.4	1972	US-08-875-435B-4	Sequence 4, Appl
29	288.5	4.4	2468	US-09-976-594-726	Sequence 726, Appl
30	288.5	4.4	2468	US-09-538-092-1135	Sequence 1135, Ap
31	288.5	4.4	2522	US-09-949-016-10237	Sequence 10237, A
32	288	4.4	1976	US-09-538-092-1078	Sequence 1078, Ap
33	286	4.3	1162	US-08-728-323A-2	Sequence 2, Appl
34	286	4.3	1162	US-09-298-568-2	Sequence 2, Appl
35	286	4.3	1162	US-09-410-399-2	Sequence 2, Appl
36	286	4.3	1162	US-09-894-273-2	Sequence 2, Appl
37	285.5	4.3	2482	US-08-328-254-6	Sequence 6, Appl
38	284.5	4.3	3210	US-09-538-092-1154	Sequence 1154, Ap
39	284.5	4.3	3248	US-08-353-700-1	Sequence 1, Appl
40	284.5	4.3	3248	PCT-US95-16216-1	Sequence 1, Appl
41	283.5	4.3	1972	US-08-875-435B-3	Sequence 3, Appl
42	280.5	4.2	1969	US-09-418-710-72	Sequence 72, Appl
43	280.5	4.2	1969	US-09-839-479-71	Sequence 71, Appl
44	280	4.2	1663	US-08-603-753D-2	Sequence 2, Appl
45	280	4.2	1663	US-09-099-753-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-949-016-10326
; Sequence 10326, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10326
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10326

Query Match      36.7%; Score 2424; DB 4; Length 1089;
Best Local Similarity 48.9%; Pred. No. 1,1e-162;
Matches 553; Conservative 165; Mismatches 234; Indels 178; Gaps 34;

231 EAIRAKMMSKLS--HSAENFEDFELDTGNGESP-EKKRERKAARLGKAMKQMSFTQ 287
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1 KSTRQKKEPSLESGVHSE-----EGSELSTKGTTRKERRKAARLSKALQKLSFTQ 51

288 RIIRSSVSPLHLEPPTIHDFFKRRPRPLCOGNAMOLIKSTKQPC----- 335
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
52 RIIRSSALNLPHPMBENKTIHDFKRRPRPTCHGNAMALLKSSKYQSSHHKEIIDTANT 111

336 ---TEKKKPNELICAEVPEF-----DYSKEDLEISPEOP 368
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
112 ENNSHSHKSGSQTTGANEVETNALPVYSKETQIITGSDSCRDLVKNBELFQEKOK 171

369 LINTQCSHAVALCVQNDARTGLSKSTEAIVVTGGMNDHEDAFSDSNIVHEQETVGLTV 428
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
172 -----QSDIRP-----SPGDSVYLQGRSN-----FLGNHSEECQVGLVAF 208

429 TETPQPTPIPOESVYQGIQI--DVEVMQMPQ-----PTHKRKLKLEKLKALG 478
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
209 -----EHALEGGEPQPEETDEKVEPEQONKSAVGPKEKRRRTLDLKKOLG 258
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
479 VLSIKPRLCPDDGSFVNLDEPKPKPEALKEPLKATLQKSKPRTKRYNLNIIRKET 538
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:42:31 : Search time 156.144 Seconds
(without alignment)

4214.186 Million cell updates/sec

Title: US-09-982-091a-2

Perfect score: 6604

Sequence: 1 MAALCEEQVLEPEPDISLK.....RFRKSTPTVKSRSIFQLE 1285

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: UniProt_03:

1: uniprot_sprot:

2: uniprot_cprot:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	6604	100.0	Q9DF50	Q9DF50 xenopus lae
2	2795	42.3	Q80YR7	Q80YR7 mus musculi
3	2687	40.7	Q9H4M4	Q9H4M4 homo sapien
4	1101	16.7	Q8JGM2	Q8JGM2 mus musculi
5	1098.5	16.6	Q8JW11	Q8JW11 homo sapien
6	861.5	13.0	Q8JX02	Q8JX02 homo sapien
7	786	11.9	Q8JRB5	Q8JRB5 dirosophila
8	722	10.9	Q8JX01	Q8JX01 homo sapien
9	579.5	8.8	Q7PR25	Q7PR25 anopheles g
10	496.5	7.5	Q8JYR6	Q8JYR6 dirosophila
11	462.5	7.0	Q7J838	Q7J838 caenorhabdi
12	421.5	6.4	Q8JRB3	Q8JRB3 plasmodium
13	420	6.4	Q8J7F7	Q8J7F7 neurospora
14	416.5	6.3	Q8JH98	Q8JH98 dirosophila
15	410	6.2	Q7RTL4	Q7RTL4 dirosophila
16	409	6.2	Q8J7G8	Q8J7G8 dirosophila
17	409	6.2	Q8J7G8	Q8J7G8 dirosophila
18	408	6.2	Q8J7G8	Q8J7G8 dirosophila
19	404	6.1	Q8J7G8	Q8J7G8 dirosophila
20	396	6.0	Q8J7G8	Q8J7G8 dirosophila
21	393.5	6.0	Q8J7G8	Q8J7G8 dirosophila
22	392	5.9	Q8J7G8	Q8J7G8 dirosophila
23	390.5	5.9	Q8J7G8	Q8J7G8 dirosophila
24	390	5.9	Q8J7G8	Q8J7G8 dirosophila
25	389.5	5.9	Q8J7G8	Q8J7G8 dirosophila
26	389	5.9	Q8J7G8	Q8J7G8 dirosophila
27	386.5	5.9	Q8J7G8	Q8J7G8 dirosophila
28	380	5.8	Q8J7G8	Q8J7G8 dirosophila
29	378.5	5.7	Q8J7G8	Q8J7G8 dirosophila
30	375.5	5.7	Q8J7G8	Q8J7G8 dirosophila
31	373	5.6	Q8J7G8	Q8J7G8 dirosophila

32	373	5.6	2762	2	Q9VGW4	Q9VGW4 dirosophila
33	370	5.6	2492	1	ATRX_PANTR	ATRX_PANTR pan troglod
34	368	5.6	18519	2	Q8JSE6	Q8JSE6 caenorhabdi
35	368	5.6	18534	2	Q8JSE7	Q8JSE7 caenorhabdi
36	367	5.6	5412	1	FUTS_DROME	FUTS_DROME drosophila
37	365	5.5	2492	1	ATRX_PONPY	ATRX_PONPY pongo pygma
38	364.5	5.5	1481	2	Q8JVB9	Q8JVB9 mycoplasma
39	364.5	5.5	2116	1	MYS2_DICDI	MYS2_DICDI dictyosteli
40	362	5.5	1578	2	Q8JSE5	Q8JSE5 caenorhabdi
41	359.5	5.4	1110	2	Q9J2S5	Q9J2S5 petromyzon
42	358	5.4	1871	2	Q8JRD5	Q8JRD5 arabidopsis
43	352	5.3	1596	2	Q8J1J4	Q8J1J4 plasmodium
44	351.5	5.3	1430	2	Q8JG29	Q8JG29 dirosophila
45	350	5.3	1431	2	Q96T23	Q96T23 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9DF50	PRELIMINARY:	PRT:	1285 AA.
AC	Q9DF50			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Caaspin.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;			
OC	Xenopodidae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=2100493; PubMed=11090622;			
RA	Kumagai A., Dunphy W.G.;			
RT	"Caaspin, a novel protein required for the activation of Chk1 during a			
RT	DNA replication checkpoint response in Xenopus egg extracts.";			
RL	Mol. Cell 6:839-849(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Kumagai A., Dunphy W.G.;			
RL	Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AF297867; AAC24516.2; -			
SQ	SEQUENCE 1285 AA; 145981 MW; A302479768DBA7D CRC64;			

Query Match 100.0%; Score 6604; DB 2; Length 1285;
Best Local Similarity 100.0%; Pred. No. 1.2e-234;
Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAALCEEQVLEPEPDISLKIVETDSGSGSCEMADONKILGCVEDKDTDEILVRKKS	60
DB	1	MAALCEEQVLEPEPDISLKIVETDSGSGSCEMADONKILGCVEDKDTDEILVRKKS	60
QY	61	KKKEVLVDSDEELKEMNPAVDNVKGSNDNEETNATYREKPKIRSAVSDNSDHE	120
DB	61	KKKEVLVDSDEELKEMNPAVDNVKGSNDNEETNATYREKPKIRSAVSDNSDHE	120
QY	121	LDVQISTQNAAEIPESHDLSLEKETHVTKPKTSKSLKQDTNKEEIVKSKKRIPE	180
DB	121	LDVQISTQNAAEIPESHDLSLEKETHVTKPKTSKSLKQDTNKEEIVKSKKRIPE	180
QY	181	KIKRTKQSKAVAEARNLNDSCGLTLDGDLFNGVNEVNDNSNEEESLAIIRAKMSK	240
DB	181	KIKRTKQSKAVAEARNLNDSCGLTLDGDLFNGVNEVNDNSNEEESLAIIRAKMSK	240
QY	241	LNLSAENFEDFEDTDEGNQSPKPKRKARLGEKMKMHSBETORLIRESSVSLPYH	300
DB	241	LNLSAENFEDFEDTDEGNQSPKPKRKARLGEKMKMHSBETORLIRESSVSLPYH	300
QY	301	LPEPKTHDFPKRPRPLCCGNAMOLIKSTYQPCTEEKKKPNEIEICAEVDFPVSKED	360
DB	301	LPEPKTHDFPKRPRPLCCGNAMOLIKSTYQPCTEEKKKPNEIEICAEVDFPVSKED	360

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:53:46 : Search time 36.3355 seconds
(without alignments)
3402.693 Million cell updates/sec

Title: US-09-982-091A-2
Perfect score: 6604
Sequence: 1 MAALCEERQVLEPEDISLK.....RFRKDSTPVKSRISIFQLE 1285

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	462.5	7.0	725	2 T21363	hypothetical prote
2	392	5.9	1526	2 A45605	mature-parasite-in
3	364.5	5.5	2116	2 A26655	myosin heavy chain
4	360	5.5	3488	2 T34418	hypothetical prote
5	359.5	5.4	1110	2 I51116	NF-180 - sea lamp
6	358	5.4	1871	2 D66796	probable heat shoc
7	355.5	5.4	5327	2 T13564	microtubule-associ
8	341.5	5.2	1909	2 A45592	liver stage antige
9	340.5	5.2	1658	2 S55101	hypothetical prote
10	336.5	5.1	1898	1 A45973	trichomyalin - hum
11	328.5	5.0	2447	2 T16870	hypothetical prote
12	323	4.9	1624	2 T25592	hypothetical prote
13	322.5	4.9	1922	2 T00637	hypothetical prote
14	322	4.9	1271	2 A45555	glutamate rich pro
15	321.5	4.9	3225	2 I52300	giantin - human
16	321.5	4.9	3259	1 A56539	giantin - human
17	320.5	4.9	1558	2 B71603	RESA-H3 antigen PF
18	320.5	4.9	1875	2 S38173	myosin-like protei
19	318.5	4.8	678	2 A54514	glutamic acid-rich
20	317.5	4.8	2020	2 T21174	hypothetical prote
21	315.5	4.8	1621	2 A82255	hypothetical prote
22	315.5	4.8	1790	2 S67593	transport protein
23	315	4.8	1744	2 JH0720	canabin - African
24	313	4.7	1979	2 C71622	hypothetical prote
25	312	4.7	2139	2 T18296	myosin heavy chain
26	311.5	4.7	1019	2 T50251	hypothetical prote
27	310	4.7	1233	2 S56271	hypothetical coile
28	309	4.7	1877	2 T21861	hypothetical prote
29	307.5	4.7	2364	2 A56577	microtubule-associ

30	307	4.6	1690	2 T13030	microtubule bindin
31	306	4.6	1939	2 T18372	repeat organellar
32	305	4.6	2484	2 T26216	hypothetical prote
33	305	4.6	2607	2 T26215	hypothetical prote
34	304	4.6	2464	1 QRMSPI	microtubule-associ
35	303	4.6	5170	2 T15348	hypothetical prote
36	302.5	4.6	1999	1 S21801	myosin heavy chain
37	302.5	4.6	2738	2 E88320	protein P07A11.6
38	300	4.5	1620	2 S61535	nucleotide-binding
39	300	4.5	1827	2 T16270	hypothetical prote
40	299.5	4.5	2954	2 T14156	kinesin-related pr
41	298.5	4.5	1992	2 A47297	myosin heavy chain
42	297.5	4.5	2722	2 T20532	hypothetical prote
43	295	4.5	1940	1 S04090	myosin heavy chain
44	294.5	4.5	1927	2 A59236	embryonic muscle m
45	294	4.5	2253	2 T30336	nuclear/mitotic ap

ALIGNMENTS

RESULT 1

T21363
hypothetical protein F25H5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21363
R:Steward, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: 219412
A:Accession: T21363
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-725 <W11>
A:Cross-references: UNIPROT:O17838; EMBL:Z81068; PIDN:CAB02986.1; GSPDB:GN00019; CESP:F21
A:Experimental source: clone F25H5
A:Gene: CESP:F25H5.5
A:Map position: 1
A:Intons: 5/3; 108/1; 129/1; 233/2; 294/2; 529/2; 601/3

Query Match 7.0%; Score 462.5; DB 2; Length 725;
Best Local Similarity 24.8%; Pred. No. 8.1e-12;
Matches 210; Conservative 122; Mismatches 283; Indels 233; Gaps 36;

QY	448	IQNDVVEQRMPEQPTHKPKLSKEKALGVDSIKP----	RLCPDDGSFVNL--DE	499
DB	9	ISSENTEVDKQAEEDR--KLRRLRRKALALERSPLPVPKLGRLVLSANSTIDLTNEDD	65	
QY	500	PKPKNEFEALKEKRFKHTLQSKPRTERKVNINIRKETTADGKEELKADVVPIMATK	559	
DB	66	DPSNKKWF-----KTFNAEKIRKLASPEAKLPVAPRKISKNA	104	
QY	560	PDKSIYQKPEKTLQVLKVLQEQAMKIRSEERLTKQALYKLDNE--GFEDDEEEEMT	616	
DB	105	P-KSLITS-----ETLKKKQLQASMAKRRKQEQRKMYEEDNEHLKKGSDEEEDVVF	158	
QY	617	EESEDDGQNAETADYPCGDEBEVEYGAEDDNDDEDITVNDRLGAVPEIVLPPTVTD	676	
DB	159	KKKKKKKEVNEEYKSEDEDEYEPENSEDESD-----	192	
QY	677	SLMLFKNSSKLGSLDDESGCKSSRLFEYEDBLDPOLKENSHNSFELISSMIPSYQ	736	
DB	193	-----KNGRAD-----EMSVNLINDSF-----TYD	214	
QY	737	PCNKTTRVINSNMLGFRSPSPVHKTSFLSSASKSGKMSSESP--VEDSQDLYNASP	794	
DB	215	VFNWVKRSGPGSTVSNFDSRSE-----ADVVEKNGVQLPTWTEITGSSNDILNCP	266	
QY	795	EPKASLYLCAGNSQGFLEDDT-----QSLLDADGLNAGR	832	
DB	267	GSFGVSQAGPPDITLQMLEIDTTERKINLPONSTVSVDKPRLESESEDEDD-VGVKR	325	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:40:50 : Search time 137.977 Seconds
(without alignments)
3601.964 Million cell updates/sec

Title: US-09-982-091A-2
Perfect score: 6604

Sequence: 1 MAALCEEQVLEPPEDISLK.....RFRKSTPTVKSRSIFQLE 1285

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_GeneSeq1Dec04:*

1: geneseqp1980a:*\n2: geneseqp1990a:*\n3: geneseqp2000a:*\n4: geneseqp2001a:*\n5: geneseqp2002a:*\n6: geneseqp2003a:*\n7: geneseqp2003bs:*\n8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6604	100.0	1385	5	AAU97586 Xenopus C
2	2681	40.7	1332	7	ADG31665 Human nov
3	2687	40.7	1332	5	AAU97587 Human C1A
4	784.5	11.9	1812	4	ABBS8022 Drosophila
5	473.5	7.2	191	5	ABBS9891 Human mac
6	390.5	5.9	681.5	4	ABBS6811 Drosophila
7	378.5	5.7	3257	4	ABBS67502 Drosophila
8	373	5.6	2482	8	ADG98534 Human ATR
9	372	5.6	2375	7	ADG68471 Human hea
10	372	5.6	2492	8	ADG60200 Angiogene
11	372	5.6	2492	8	ADG60501 Angiogene
12	372	5.6	2492	8	ADP73124 Angiogene
13	371	5.6	2288	8	ADG60224 Angiogene
14	371	5.6	2288	8	ADG60525 Angiogene
15	371	5.6	2288	8	ADP73148 Angiogene
16	365.5	5.5	1234	8	ADP98957 C. albica
17	354	5.4	1398	5	ABG97491 Human NOV
18	354	5.4	1400	4	ABBS1348 AAP-2 pro
19	353	5.3	1445	6	ABR41365 Human DIR
20	341	5.2	2887	4	ABBS8514 Drosophila
21	336.5	5.1	1898	2	ABY30795 A human c
22	336.5	5.1	1898	7	ADG48869 Human pro
23	333.5	5.0	2230	6	ABU07445 Protein d
24	333.5	5.0	2230	7	ABR61600 Human gol
25	333.5	5.0	2230	7	ABR61600 Human gol

ALIGNMENTS

26	333.5	5.0	2250	7	ABR61601	ABR61601 Human gol
27	333.5	5.0	2252	7	ABR61602	ABR61602 Human gol
28	331.5	5.0	3111	4	ABBS60327	ABBS60327 Drosophila
29	328.5	5.0	2476	7	ADP1512	ADP1512 Murine AT
30	327.5	5.0	1584	8	ABM80133	ABM80133 Tumour-as
31	324	4.9	1881	5	ABP73809	ABP73809 Candida a
32	321.5	4.9	3225	7	ADJ68448	ADJ68448 Human hea
33	321.5	4.9	3259	7	ADBS6037	ADBS6037 Human pro
34	321.5	4.9	3259	7	ADBS6033	ADBS6033 Human pro
35	320.5	4.9	1558	3	AAH18924	AAH18924 Plasmodiu
36	320.5	4.9	1875	6	ABR53560	ABR53560 Protein s
37	320.5	4.9	1875	7	ADK64380	ADK64380 Disease c
38	320.5	4.9	1875	8	ADG43855	ADG43855 Bacteri
39	318	4.8	1690	4	ABBS61144	ABBS61144 Drosophila
40	318	4.8	1690	4	ABBS61173	ABBS61173 Drosophila
41	317.5	4.8	2020	8	ADN22475	ADN22475 Bacteri
42	317.5	4.8	2020	8	ADN22476	ADN22476 Bacteri
43	315.5	4.8	1790	6	ABR53116	ABR53116 Protein s
44	315.5	4.8	1790	7	ADK63040	ADK63040 Disease c
45	315.5	4.8	1790	8	ADN18933	ADN18933 Bacteri

RESULT 1

AAU97586 standard; protein; 1285 AA.

AC	AAU97586;	Location/Qualifiers
DT	27-AUG-2002 (first entry)	
XX	Xenopus Claspin protein.	
DE	Xenopus Claspin protein.	
XX	Chk1 protein; SQ/TQ motif; isoelectric point; cell cycle progression;	
KW	nuclear localisation signal; DNA replication checkpoint; benign neoplasm;	
KW	cell proliferative disorder; malignant neoplasm; frog; claspin.	
XX	Xenopus sp.	
OS	Xenopus sp.	
XX		
FT	Key	Location/Qualifiers
FT	Region	158..174
FT		/note="Nuclear localisation signal site"
FT	Region	312..316
FT		/note="Nuclear localisation signal site"
FT	Region	1078..1084
FT		/note="Nuclear localisation signal site"
XX		
PN	WO200233115-A2.	
XX		
PD	25-APR-2002.	
XX		
PF	17-OCT-2001; 2001WO-US032316.	
XX		
PR	17-OCT-2000; 2000US-0241246P.	
XX		
PA	(CALY) CALIFORNIA INST OF TECHNOLOGY.	
PI	Kumagai A, Dunphy WC;	
XX		
DR	WPI: 2002-454610/48.	
XX	N-PSDB; ABR52610.	
XX		
PT	Novel Claspin polypeptide specifically interacting with chk1 protein	
PT	useful for identifying compound that modulates cell cycle progression and	
PT	for treating cell proliferative disorder like neoplasm.	
XX		
PS	Claim 2; Fig 1; 97pp; English.	
XX		
XX	The present invention relates to a new substantially pure Claspin	
CC	polypeptide that specifically interacts with a Chk1 protein, having SQ/TQ	
CC	motifs, an isoelectric point of 4.5 and at least one nuclear localisation	